

Perspective

Large-scale molecular characterization and analysis of gastric cancer

Ilya Shmulevich

The recent effort by The Cancer Genome Atlas (TCGA) Network has revealed that gastric cancer, which is a leading cause of cancer-related deaths worldwide with a 5-year survival rate less than 25%, is a much more heterogeneous disease than previously thought. And yet, conventional treatment approaches and clinical trials have assumed it is a single disease. Although it is well known that under the microscope, gastric cancer cells appear quite different, the current classification scheme recognizes two main categories of gastric cancer: diffuse and intestinal.

By integrating large-scale measurements of DNA, RNA, and proteins from 295 primary gastric tumors, the TCGA project found that there are four major subtypes of gastric cancer. The fact that these subtypes appear quite distinct on the molecular level suggests not only that the molecular processes driving tumorigenesis can vary among patients but also that the treatments may have to be tailored based on the subtype of tumor a patient has. This molecular stratification of patients, enabled by large-scale molecular characterization, is a significant step towards personalized therapy.

The data that are generated from tumors in the TCGA project are not only very large in terms of volume, but, perhaps even more importantly, highly heterogeneous. Six different molecular analysis technologies were used for the molecular characterization of gastric cancer. These technologies provided information on DNA mutations, amplifications or deletions of gene sequences, epigenetic modifications of DNA, and levels of mRNA, microRNA, and proteins. Additionally, coded clinical information on pathology, histology, tumor characteristics, and other relevant data on each patient was collected.

The computational challenge is how to integrate all this information not only to see broad differences among tumors but also to identify strong statistical associations among all these molecular and clinical data. Such associations may provide clues to how molecular systems in cancer cells are disrupted in different subtypes of gastric cancer and what treatment strategies may be most effective.

For example, one of the four subtypes is characterized by the

presence of the Epstein-Barr virus (EBV). Patients having tumors of this subtype tend to also have mutations in the PIK3CA pathway, extreme DNA hypermethylation, and extra copies of *PD-L1* and *PD-L2* genes, which are suppressors of immune response. These findings suggest that inhibitors of the PI3-K pathway may be of potential use for this subtype of gastric cancer. Further, *PD-L1/2* antagonists may help promote immune destruction of the tumor.

Another example of a statistical association that may have clinical use is the finding that frequent mutations in a gene called *RHOA* occur predominantly in the subtype of gastric cancer termed “genomically stable,” which is characterized by the lack of high levels of aneuploidy and are predominantly diffuse-type tumors. These subtypes of tumors have a higher likelihood of metastasis and are more invasive. It is known that the gene product of *RHOA* interacts with other proteins to alter cellular shape and motility, which are important for tumor growth. This protein or pathway may be an important therapeutic target for the genomically stable subtype of gastric cancer.

The large-scale statistical analysis of the data revealed numerous other prevalent associations among factors, such as age at diagnosis, anatomic region of the stomach where the tumor occurs, gender, mutation status of genes such as *TP53*, and other tumor or patient characteristics. Although identifying such relationships in heterogeneous cancer data is important for the initial discovery process, much work lies ahead to functionally characterize the molecular aberrations in cancer and to develop and test effective targeted therapies.

The ever-increasing volume and complexity of molecular data from cancer, exemplified by the TCGA project, calls for more sophisticated analytical approaches that can detect multivariate relationships in heterogeneous data and map that information onto the existing body of knowledge of the structure of molecular networks that govern all cellular processes. Other large-scale molecular characterization projects, such as the ENCyclopedia Of DNA Elements (ENCODE) and the Genomics of Drug Sensitivity in Cancer, are producing highly complementary data that will allow scientists to paint a more complete picture of how molecular regulatory processes are disrupted in cancer cells and what drugs may be effective in targeting a cancer’s weaknesses. This first large-scale molecular characterization and analysis of gastric cancer undertaken by the

Author's Affiliation: ISB/MDACC Genome Data Analysis Center, TCGA Gastric Cancer Working Group, Institute for Systems Biology, Seattle, WA 98109, USA. Email: ilya.shmulevich@gmail.com.

doi: 10.5732/cjc.014.10116

TCGA project takes us one step closer toward helping patients

suffering from this deadly type of cancer.

References

[1] The Cancer Genome Atlas Research Network. Comprehensive molecular characterization of gastric adenocarcinoma. Nature,

2014. DOI: doi:10.1038/nature13480.



The poster is a promotional graphic for the 2014 World Cancer Congress. At the top, it features three logos: a stylized red and orange circular logo on the left, the text '2014 | DEC 3-6 MELBOURNE AUSTRALIA' in the center, and the slogan 'JOINING FORCES – ACCELERATING PROGRESS' on the right. To the far right is the UICC logo, which includes a globe icon and the text 'uicc global cancer control WORLD CANCER CONGRESS'. The main body of the poster has a light beige background with a large, colorful, abstract graphic on the right side consisting of concentric arcs and dots in shades of yellow, orange, pink, and purple. The text 'SAVE THE DATE' is written in large, bold, red capital letters. Below it, in blue capital letters, is 'IF THERE IS ONE EVENT YOU NEED TO ATTEND IN 2014, IT'S THE WORLD CANCER CONGRESS'. The dates '3-6 DECEMBER 2014' are prominently displayed in large, bold, red capital letters, underlined with a thick red line. Below this, in smaller blue capital letters, is 'PRELIMINARY PROGRAMME NOW AVAILABLE'. Further down, in blue capital letters, is 'PLEASE VISIT', followed by the website 'WORLD Cancer CONGRESS.ORG' in bold red capital letters. At the bottom, there are three logos: 'MELBOURNE VICTORIA AUSTRALIA' on the left, 'Hosted by' in the center, and the 'Cancer Council Australia' logo on the right. At the very bottom, in small grey text, it says 'Held in conjunction with the Clinical Oncology Society of Australia (COSA) Annual Scientific Meeting'.

2014 | DEC 3-6
MELBOURNE AUSTRALIA

**JOINING FORCES –
ACCELERATING
PROGRESS**

uicc
global cancer control
WORLD CANCER
CONGRESS

SAVE THE DATE

**IF THERE IS ONE
EVENT YOU NEED
TO ATTEND IN 2014,
IT'S THE WORLD CANCER
CONGRESS**

**3-6 DECEMBER
2014**

PRELIMINARY PROGRAMME NOW AVAILABLE

**PLEASE VISIT
WORLD Cancer CONGRESS.ORG**

MELBOURNE
VICTORIA AUSTRALIA

Hosted by

**Cancer Council
Australia**

Held in conjunction with the Clinical Oncology Society of Australia (COSA) Annual Scientific Meeting